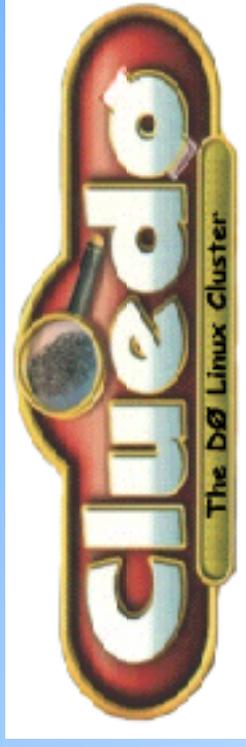


ClueD0 - The Sequel



- Current cluster
- Why change?
- Overview of changes
- Cluster Splitting
- Kai and other problems/hacks
- Misc. - printing
- RH7.1 vs. ClueD0 Linux 7.1
- CLuMP
- Security
- Batch - PBS

Current ClueD0

- 59 machines (at least 17 waiting) - probably top 100 this summer
- 130 users, 19 institutes
- 7 administrators of varying levels of involvement
- Based on RedHat 6.1 (with many upgrades)
- Provide
 - home directories + backup (actually, you do that)
 - D0 code distribution (ok, you do that too)
 - security - kerberos, tripwire, patches
 - web server
 - batch queues
 - printing
 - friendly advice

Why Change??

- **Our current operating system is old**
- **new machines become difficult to install**
- **flaws in old software become more difficult to live with (eg. NFS)**
- **user environment becomes antiquated (or we spend time updating)**
- **We did not originally intend to scale the cluster to 100s of machines**
- **our administrative burden needs to be split more evenly**
- **the server burden needs to be split**
- **we need to ensure robustness (multiple servers)**

Overview of Changes

- Split cluster into 4 (or more) sub-clusters
- Upgrade to linux distribution based on RH7.1
- Separate login node, web server from cluster server
- Create aliases (eg. `www-clued0`, `plum-clued0`, etc.) for important services
- Move from NIS to LDAP for configuration info. Use CLuMP.
- Change from PILOT.FNAL.GOV realm to FNAL.GOV
- Change batch from NQS to PBS

Cluster Splitting

- We are currently a single cluster with one server, one set of administrators
- What does split mean?
 - divide cluster into logical groups of 20-30 machines
 - assign non-overlapping groups of sys-admins to each
 - each subcluster has its own server (robustness)
- We mount homedirs and d0dist from d02ka. This makes splitting the cluster very easy. Subcluster servers only need to act as LDAP slave servers and provide rsync'ed copy of /usr/local/
- No difference in environment, batch queues, etc. between subclusters.
- Try to make split based on geography...one subcluster per building to start

Splitting (cont.)

- **What does this require?**
 - institutes with several machines provide a sys-admin for their sub-cluster
 - an institute in each cluster provides a machine to act as "server"
- **What does an admin do?**
 - install the machines in their subcluster
 - first in line to answer questions from their users
 - keep their server up

Splitting.....Server Stuff

- In addition to subcluster servers (LDAP slave + /usr/local - gigabit?!) we will have several disk servers
- So far, 4 disk servers (640Gb each from ASA Computers) all in PK173
- We expect each subcluster to eventually have at least one disk server with significant shared space. These should have gigabit connections.
- One of our existing disk servers will become a SAM station...gigabit for sure.
- So, we could use at least one gigabit port per building, but expect this need to grow quickly.

Kai and Other Hacks

- We tested compiling, linking and running D0 code on a RH7.1 system
- UPS/UPD needed a little fiddle because our kernel is 2.4, not 2.2
 - `cat /fnal/ups/etc/ups_override.ripon-clued0 -H Linux+2.2.44` (44 is just a note-to-self)
- NFS mount of /d0dist from d02ka defaults to NFS v3. This should be a good thing....
 - `in /d0dist/dist/lib/Linux_blah/` I see lots of files....but not all files
 - forcing mount to use NFS v2 allows me to see all files
- KCC v3.4 does not (and will not) support RH7.1. We currently use 3.4 to compile the D0 code. KCC v4.0e does support RH7.1

Hacks...Kai....

- **Ultimate solution: upgrade to KCC 4.0e- lots of work**
- **Temporary solution: hack KCC v3.4 and cross fingers**
- **The hack:**
 - **steal c header files from 6.1 system (tar /usr/include)**
 - **untar into /usr/old-include on 7.1 system**
 - **edit files in**
`/fnal/ups/kai/v3_4gnode_locked/KCC_BASE/include`
`/fnal/ups/kai/v3_4gnode_locked/KCC_BASE/include/sys`
replacing /usr/include with /usr/old-include: `errno.h`,
`math.h`, `pthread.h`, `stdlib.h`, `wchar.h`, `sys/cdefs.h`, `sys/stat.h`,
`sys/types.h`.
 - **link /usr/old-include/linux to /usr/include/linux and**
`/usr/old-include asm to /usr/include/asm`

Hacks....Kai....

- **To compile code**
 - `setenv USR_INCLUDE /usr/old-include`
 - `setup D0RunII p08.11.00`
 - `newrel, addpkg, gmake, etc.`
- **Tests/Results**
 - **pre-linked executables from 6.1 build machines run without problem on a 7.1 machine (as expected)**
 - **linking executables based on 6.1-only libraries works without problem (d0reco passed tests, d0trigsim produces identical results)**
 - **Compiling some packages under 7.1 and linking with 6.1 pre-compiled packages produces good executable. Some warnings which appear to be spurious:**
`/fnal/ups/kai/v3_4gnode_locked/KCC_BASE/bin/nm:
/usr/lib/libpthread.so: no symbols`

Hacks...Kai...Tests

- executable compiled/linked on 7.1 system does not run on 6.1 system (GLIBC_2.2 not found)...as expected.
- attempted to compile a whole release (t01.46.00) from scratch on 7.1 system. Failed in external packages. eg.

CLHEP:

```
<--> (SRT_rules.mk; __CLHEP__) Compiling static library  
object BoostX. "/usr/include/sys/select.h", line 99: error  
#130: expected a "{" struct timeval *__restrict __timeout)  
__THROW;
```

ISO Cxx:

```
=====  
Probe for general library defects: =====  
checking abs(double) via abs_dbl.cc... defective  
checking cure for abs(double) defect via abs_dbl.cc...  
PROBLEM! ....etc.
```

Hacks...Kai...Tests

- **try instead to check out every package (~490) and compile it (as a user would) and compare failures to 6.1 failures on same release. 2 (min) errors on 7.1 not present in 6.1**

```
<--> Compiling static library D0OM linkage for component
BaseGeometry_Ink.
"/disk2/work/oneil/t146/include/d0om/private/d0_Ref.ipp",
line 332: error #70:
incomplete type is not allowed
(d0_Ref_Base::ptr1 (T::d0om_type_static ()))
^
detected during instantiation of
"T *d0_Ref<T>::ptr1() const [with T=SiBaseGeometry]"
```

Hacks...Kai...Tests

```
<--> Compiling static library object Callback.  
"/usr/include/bits/sched.h", line 57: error #130: expected a "{"  
int __flags, void * __arg) __THROW;  
Error limit reached.  
100 errors detected in the compilation of  
"/disk2/work/oneil/t146/framework/src/Basic.cpp".
```

Hacks...Kai...Conclusions

- We should make a timetable for moving to KCC 4.0e. Fermilinux will soon see the same problems we see...
- We have a hack to make KCC 3.4g work most of the time. We can compile D0-specific packages...it is possible that users won't even notice any difference. More investigation even fix all packages to work with the hack...but someone would have to be willing to do that....
- Executables from 6.1 run on 7.1 and linking against existing 6.1 D0 libraries works fine.

Misc. - Printing

- Just in case you want to know how we manage printing....
- Common Unix Printing System (CUPS) and ESP PrintPro.
 - One server license (\$250 for linux) and infinite number of clients. All clients read config from server (add printer once)
 - gui print manager and html-based manager (localhost:631)
 - gui print interface for users (glp)
 - commercial printer drivers
 - anyone with cups installed (free) can walk up (laptop) and see all printers broadcast by servers
- <http://www.easysw.com/printpro/>

